

# ESTADISTICA APLICADA A LA INVESTIGACION CIENTIFICA

## Introducción al Diseño de Experimentos (DOE)

# Breve reseña histórica

Ronald A. Fisher (1900-1950) en Inglaterra, utilizó para conseguir un aprendizaje acelerado. (The Design of Experiments, 1935). **Investigación agrícola.**



George E. P. Box desarrolló la metodología de superficie de respuestas (Box y Wilson, 1951). (**Industria química**). Incluyó nuevas familias de diseños y una estrategia para la experimentación secuencial. Entre 1950 y 1980, el DOE se convirtió en una herramienta de aplicación frecuente, pero sólo en las áreas de investigación y desarrollo.



En la década de 1980 se dio un gran impulso debido al éxito en **calidad de la industria** japonesa. Promovió el uso de la estadística en calidad. En Japón destaca el trabajo de Genichi Taguchi, cuyos conceptos sobre diseño robusto también tuvieron un impacto significativo en la academia en el mundo occidental.



# Definición

Metodología estadística que sirve para ***planificar, ejecutar y analizar experimentos de manera organizada*** con el objetivo de entender como diferentes factores influyen en un resultado y encontrar la mejor combinación de esos factores para optimizar un proceso, producto o servicio.

Es la aplicación del método científico para generar conocimiento acerca de un sistema o proceso, por medio de pruebas planeadas adecuadamente.

Consiste en la manipulación sistemática de las variables de entrada de un proceso para entender el efecto que estas pueden causar en la variable respuesta.

# Experimento

Es un cambio en las condiciones de operación de un sistema o proceso, que se hace con el objetivo de medir el efecto del cambio en una o varias propiedades del producto o resultado.

# Unidad Experimental

Pieza(s) o muestra(s) que se utiliza para generar un valor que sea representativo del resultado de la prueba.

# ¿Cuándo aplicar el Diseño de Experimentos?



- ✓ Cuando se quiere mejorar un proceso (ejemplo: aumentar la vida útil de un alimento cambiando temperatura y envasado).
- ✓ Para comparar alternativas (ejemplo: probar 3 proveedores de insumos y ver cuál da mejor calidad).
- ✓ Cuando se busca optimizar (ejemplo: encontrar la mejor combinación de tiempo y temperatura para hornear pan).
- ✓ Para entender causas y efectos (ejemplo: ver cómo influyen la velocidad y la presión en la resistencia de una pieza).
- ✓ En fase de desarrollo de productos o en investigación científica, donde se necesita conclusiones válidas y defendibles.

# ¿Cuándo **no conviene** aplicar el Diseño de Experimentos?



- ✓ Cuando no se puede controlar los factores (ejemplo: clima en agricultura a cielo abierto, donde no se puede decidir si llueve o no).
- ✓ Cuando el costo de hacer pruebas es demasiado alto y no se puede repetir ensayos.
- ✓ **Cuando los datos que se necesitan ya están disponibles (en ese caso conviene análisis estadístico observacional, no experimentos).**
- ✓ Cuando el proceso es único o irrepetible (ejemplo: un lanzamiento de cohete, donde no se puede repetir varias veces).
- ✓ Cuando no existe un método de medición confiable de la

# Datos experimentales vs. datos observacionales

**Datos experimentales:** Son los que el investigador genera controlando los factores.

**Ejemplo:** decide probar un cultivo con 2 dosis de fertilizante y 3 frecuencias de riego. Para ello se monta las parcelas, se aplica los tratamientos y se mide el rendimiento.

**“Se prueban las combinaciones de factores”  
(el investigador controla qué se cambia y cómo)**

**Datos observacionales:** Son los que ya existen, sin que el investigador lo haya planificado ni controlado.

**Ejemplo:** Un hospital ya tiene un registro de 10 años de pacientes (edad, peso, tratamientos, resultados). Aquí no se puede decidir qué pacientes recibieron X tratamiento. Solamente se observa lo que pasó. En este caso se usan técnicas de análisis estadístico observacional (regresiones, correlaciones, análisis multivariado, entre otros).

**“Se dispone de los datos, y posteriormente se analizan los datos”**

# Principios fundamentales de (DOE)

## **Aleatorización:**

Cambiar el orden de los ensayos de manera aleatoria, para que el resultado no se vea afectado por factores externos no controlados.

**Ejemplo:** Medir la resistencia de un material a dos temperaturas (200°C y 250°C) y dos tiempos (1h y 2 h). Mezclar los ensayos, los efectos externos se distribuyen y los factores son mas confiables.

## **Replicación:**

Repetir cada tratamiento varias veces. Permite estimar el error experimental y aumentar la precisión.

**Ejemplo:** Medir resistencia 3 veces a 200°C y 1h. (Luego se puede calcular promedio y desviación estándar).

## **Bloqueo:**

Agrupar unidades similares y aplicar tratamientos dentro de cada bloque. Reduce la variabilidad causada por factores no estudiados.

**Ejemplo:** Bloquear por lote de materia prima.

# Principales componentes del (DOE)

## FACTORES

VARIABLES que el investigador manipula (ej: temperatura, tiempo de cocción, dosis, entre otros).

## NIVEL

Los valores que puede tomar cada factor (ej: temperatura a 100 °C, 150°C, 200 °C).

## RESPUESTA

El resultado que se mide (ej: dureza del material, calidad del producto, rendimiento). Por lo general, estas variables se denotan con la letra “y”.

## TRATAMIENTOS

Las combinaciones de factores y niveles que se decide probar.

**ENTRADA**

**SALIDA**

Factores controlables  
Factores **NO**  
controlables



**PROCESO**

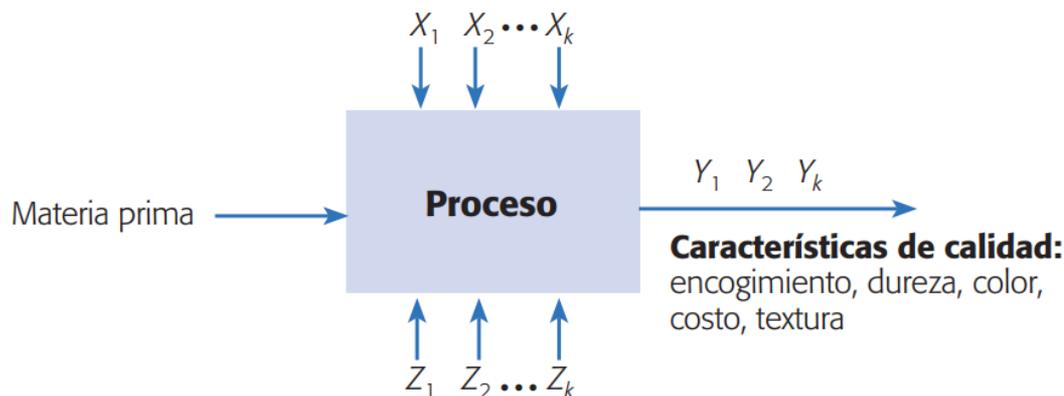


Características de  
calidad o variables  
de respuestas

Un proceso puede considerarse como una caja negra a la cual ingresan diversas variables que interactúan para producir un resultado.

**Factores de diseño (fáciles de controlar):**

Tiempo de ciclo, presión del molde, velocidad de tornillo, temperatura, tiempo de curado, contenido de humedad



**Factores de ruido (difíciles de controlar):**

- Parámetros de calidad del proveedor
- Química del plástico
- Otras variables del proceso
- Variables ambientales

Factores y variables en la fabricación de un envase de plástico (Gutierrez Pulido y Zalazar, 2008)

# Preguntas que debe realizarse el investigador



¿Cuáles son los factores controlables se van a incluir en la investigación? (tener en cuenta factores de ruidos o factores incontrolables)

¿Cuántos y que niveles involucran los factores que afectan la característica de calidad que se va a mejorar?

¿Cómo se va a medir el efecto de los factores estudiados?

¿Cuántas veces deberá ejecutarse el experimento?

¿Qué diseño experimental es el mas adecuado?

# Etapas en el diseño de experimentos

Es fundamental decidir cuáles pruebas o tratamientos se van a realizar y cuántas repeticiones de cada uno se requieren, de manera que se **obtenga la máxima información al mínimo costo posible**.

## E1) Planeación y realización

Son actividades encaminadas a entender, delimitar el problema u objeto de estudio y seleccionar variables de respuesta y factores. Concluye con la especificación de los tratamientos a realizar y con la organización del trabajo experimental.

- ✓ Entender y delimitar el problema u objeto de estudio.
- ✓ Elegir la(s) variable(s) de respuesta que será medida en cada punto del diseño y verificar que se mide de manera confiable.
- ✓ Determinar cuáles factores deben estudiarse o investigarse, de acuerdo a la supuesta influencia que tienen sobre la respuesta.
- ✓ Seleccionar los niveles de cada factor, así como el **diseño experimental adecuado a los factores** que se tienen y al objetivo del experimento.
- ✓ Planear y organizar el trabajo experimental.
- ✓ Realizar el experimento.

# Etapas en el diseño de experimentos

## E2) Análisis

Se debe recurrir a métodos estadísticos inferenciales para ver si las diferencias o efectos muestrales (experimentales) son lo suficientemente grandes para que garanticen diferencias poblacionales (o a nivel proceso). La técnica estadística central en el análisis de los experimentos es el análisis de varianza ANOVA.

## E3) Interpretación

Se debe analizar con detalle lo que ha pasado en el experimento, desde contrastar las conjeturas iniciales con los resultados del experimento, hasta observar los nuevos aprendizajes que se lograron, verificar supuestos y elegir el mejor tratamiento, con apoyo de las pruebas estadísticas.

## E4) Control y conclusiones finales

Decidir qué medidas implementar para generalizar el resultado del estudio y para garantizar que las mejoras se mantengan. Además, es difundir los logros.

# Conceptos necesarios para aplicar DOE

**Estadística básica** (medidas, hipótesis, ANOVA).

**Definición clara del problema** (factores, niveles, respuesta).

**Diseños estándar** (factorial, fraccionado, bloques, RSM).

**Planificación adecuada** (aleatorización, replicación, bloqueo).

**Análisis estadístico** (ANOVA, regresión, gráficos de interacción).

**Software de apoyo** (Minitab, R, Excel, Design-Expert).

## **Interpretación práctica para la toma de decisiones:**

Saber leer los resultados (qué factores son significativos, qué interacciones existen).  
Expresarlo con gráficos y tablas claras (por ejemplo: “la temperatura tiene más efecto que el tiempo de horneado”).

Concluir con recomendaciones prácticas para el proceso/producto

**Tipos de experimentos:** Se clasifican en función del número de factores, niveles y forma de combinación de tratamientos.



### **Diseño factorial**

prueba todas las combinaciones (confiable, pero puede ser caro).



### **Diseño factoriales fraccionados**

prueba solo parte de las combinaciones (ahorro, pero menos precisión).



### **Diseño de bloques**

controla fuentes de variación externas



### **Diseño de Superficie de Respuesta (RSM)**

busca la mejor combinación (optimización)

# Tipos de experimentos

## Diseño factorial

Dos o más factores, todos combinados en todos sus niveles. Permite estudiar interacciones entre factores).

**Diseño factorial completo ( $2^2$ )** Se prueban **todas las combinaciones posibles** de los factores.

Ensayo	Temperatura (A)	Tiempo (B)	Resistencia (MPa)
1	200°C	1 h	120
2	200°C	2 h	130
3	250°C	1 h	150
4	250°C	2 h	170

**Análisis:** La resistencia aumenta con mayor temperatura y mayor tiempo. El efecto combinado (250°C + 2h) da el mejor resultado.

# Diseño de experimento básico (factorial 2x2)

Se desea conocer la mejor combinación para preparar una buena taza de café.

## Factores:

Tiempo de infusión (Nivel: 2 min o 4 min).

Temperatura del agua (Nivel: 80°C o 90°C).

**Respuesta:** sabor (se evaluará en una escala de 1 a 10).

Se diseña un experimento probando todas las combinaciones:

2 min + 80°C

2 min + 90°C

4 min + 80°C

4 min + 90°C

Luego se analiza cuál combinación resultó con mejor sabor.



# Tipos de experimentos

## Diseño factoriales fraccionado

Si hubiera más factores (ej: presión, velocidad de enfriado, tipo de molde), el número de combinaciones crece mucho. Ejemplo: con 4 factores a 2 niveles  $\rightarrow 2^4 = 16$  ensayos.

Para ahorrar recursos, se usa un factorial fraccionado, donde se prueban solo la mitad de combinaciones (8 ensayos en lugar de 16).

**Desventaja:** Se pierde algo de información, pero se detectan los factores principales.

Usados cuando hay muchos factores y recursos limitados.  
Solo se prueban algunas combinaciones representativas.

# Tipos de experimentos

## Diseño de bloques

Factores externos (bloques) se controlan para reducir variabilidad.

Supongamos que los ensayos se hacen en dos hornos distintos (bloque: horno A y horno B). La variación entre hornos podría confundir los resultados. Para controlarlo, se organiza un diseño en bloques, donde cada combinación se prueba en cada horno.

Ensayo	Horno (Bloque)	Temperatura	Tiempo	Resistencia
1	A	250°C	1 h	118
2	A	200°C	1 h	169
3	B	200°C	1 h	122
4	B	250°C	1 h	171

El análisis separa el efecto de los factores (temperatura, tiempo) del efecto del bloque (horno).

# Tipos de experimentos

## Diseño de Superficie de Respuesta (RSM)

Cuando se quiere no solo comparar, sino optimizar, se usan diseños que permiten ajustar un modelo matemático (ej: regresión cuadrática). Ahora se prueban tres niveles de cada factor:

Temperatura: 200°C, 225°C, 250°C.

Tiempo: 1h, 1.5h, 2h.

Ensayo	Temperatura	Tiempo	Resistencia
1	200°C	1 h	118
2	200°C	2 h	132
3	250°C	1 h	152
4	250°C	2 h	171
5	225°C	1,5 h	168

Con estos datos se ajusta un modelo matemático:

$$Resistencia=80+0.3(Temperatura)+15(Tiempo)-0.05(Temperatura^2)$$

$$Resistencia=80+0.3(Temperatura)+15(Tiempo)-0.05(Temperatura^2)$$

**Interpretación:** La resistencia crece con temperatura y tiempo, pero no indefinidamente: a partir de cierto punto, subir la temperatura ya no mejora mucho. El óptimo estimado está en 225°C y 1.5h.

# Algunos ejemplos de aplicaciones

## Ciencias de la Computación (IA y Machine Learning)

Ajustar hiperparámetros de un modelo de redes neuronales.

**Aplicación del DOE:** Diseño factorial fraccionado para evaluar combinaciones de tasa de aprendizaje, número de capas y regularización.

**Beneficio:** Optimiza el rendimiento del modelo sin pruebas exhaustivas.

## Ciencias Agrarias

Determinar cómo la fertilización y el riego afectan el rendimiento del maíz.

**Aplicación del DOE:** Diseño de bloques aleatorizados con múltiples tratamientos.

**Beneficio:** Controla la variabilidad por suelos distintos y detecta combinaciones más productivas.

## Física y Materiales

Mejorar la resistencia de una aleación metálica.

**Aplicación del DOE:** Diseño Taguchi para probar diferentes proporciones de aleantes y condiciones de tratamiento térmico.

**Beneficio:** Aumenta confiabilidad del material con pocos experimentos

# OTROS EJEMPLOS

## **ingeniería Industrial**

- Diseño de líneas de producción para reducir tiempos de ciclo.
- aplicado a ergonomía: evaluar cómo iluminación, temperatura y ruido afectan el desempeño laboral.

## **Ingeniería Química**

- Optimización de procesos de polimerización considerando temperatura, catalizador y tiempo de reacción.
- Maximizar el rendimiento de biodiésel modificando proporciones de aceite, metanol y catalizador.

## **Ingeniería de Materiales**

- Determinar condiciones óptimas de sinterizado en cerámicas (temperatura, presión, tiempo).
- Estudio Taguchi para resistencia a la corrosión en aceros con diferentes composiciones de aleantes.

## **Ingeniería Eléctrica/Electrónica**

- Ajustar parámetros de soldadura en placas PCB (temperatura, tiempo de contacto, tipo de flux).
- Evaluar eficiencia energética de motores bajo variaciones de voltaje, carga y refrigeración.

## **Ingeniería Civil**

- Optimizar mezclas de concreto (cemento, arena, grava, aditivos) para mejorar resistencia a compresión.
- Evaluar impacto de diferentes asfaltos y temperaturas en la durabilidad de pavimentos.

# Caso práctico desarrollado (factorial 3X2)

## Optimización de parámetros en un proceso de torneado

### Ingeniería Mecánica

- Optimización de parámetros de torneado (velocidad de corte, avance, profundidad) para minimizar rugosidad superficial.
- Análisis de fatiga en piezas metálicas sometidas a distintas cargas y tratamientos térmicos.

**Objetivo:** Determinar cómo influyen tres factores del torneado sobre la rugosidad superficial de una pieza metálica, y encontrar la combinación óptima.

### PASO 1:

#### Definición de factores (3) y niveles(2)

- **Factor A:** Velocidad de corte (m/min) → [100, 150]
- **Factor B:** Avance (mm/rev) → [0.1, 0.2]
- **Factor C:** Profundidad de corte (mm) → [1.0, 2.0]

Usamos un **diseño factorial completo**  $2^3 = 8$  corridas experimentales.

**PASO 2:****Armado de la matriz experimental**

<b>Corrida</b>	<b>A: Velocidad</b>	<b>B: Avance</b>	<b>C: Profundidad</b>	<b>Rugosidad (Ra, <math>\mu\text{m}</math>)</b>
1	100	0.1	1.0	1.2
2	150	0.1	1.0	1.0
3	100	0.2	1.0	2.5
4	150	0.2	1.0	2.0
5	100	0.1	2.0	1.5
6	150	0.1	2.0	1.3
7	100	0.2	2.0	3.0
8	150	0.2	2.0	2.5

**PASO 3:****Análisis de efectos principales**

**Velocidad de corte (A):** A mayor velocidad, menor rugosidad (efecto positivo).

**Avance (B):** A mayor avance, la rugosidad aumenta notablemente (efecto negativo fuerte).

**Profundidad (C):** Efecto menor, pero mayor profundidad tiende a aumentar rugosidad.

#### PASO 4:

#### Detectar interacciones

**A × B:** El efecto beneficioso de la velocidad alta se pierde si el avance también es alto.

**B × C:** Avance elevado + profundidad alta = peor combinación (máxima rugosidad).

#### PASO 5:

#### Modelo estadístico (ANOVA simplificada)

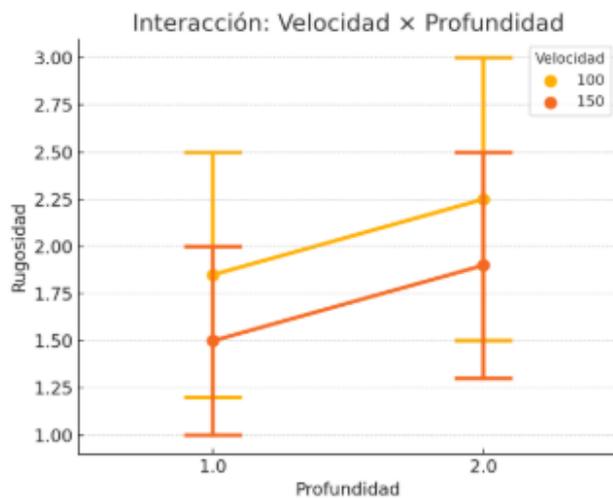
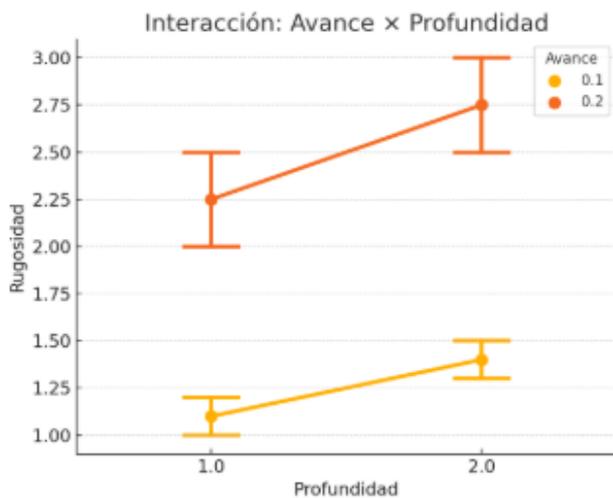
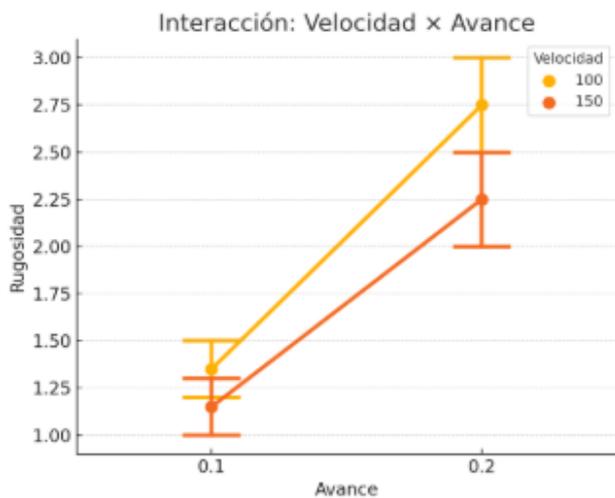
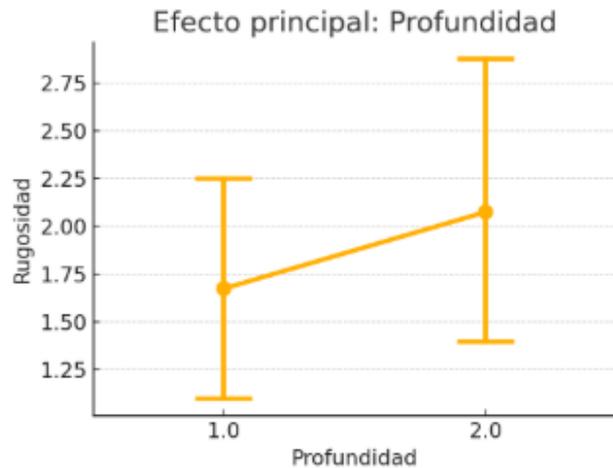
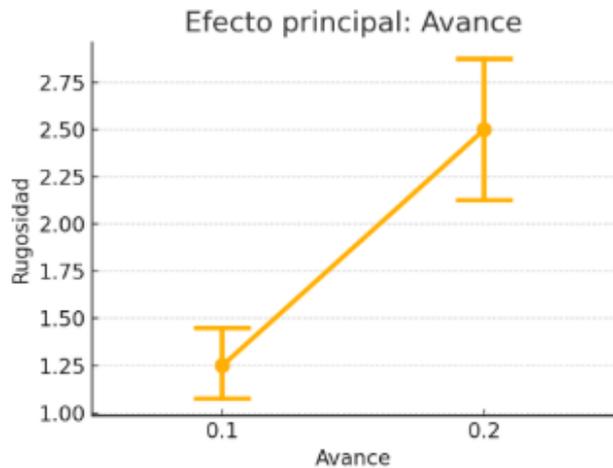
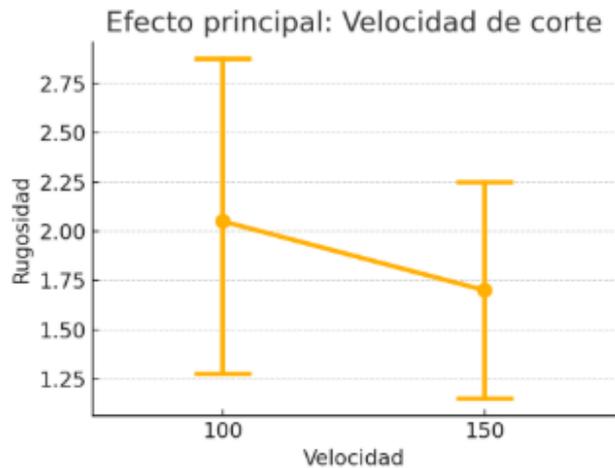
Rugosidad  $\approx \beta_0 + \beta_A(A) + \beta_B(B) + \beta_C(C) + \beta_{AB}(A \times B) + \dots$

- Resultado clave: **El avance (B) es el factor más influyente** ( $p < 0.01$ ).
- Velocidad (A) significativa pero secundaria.
- Profundidad (C) con efecto menor.

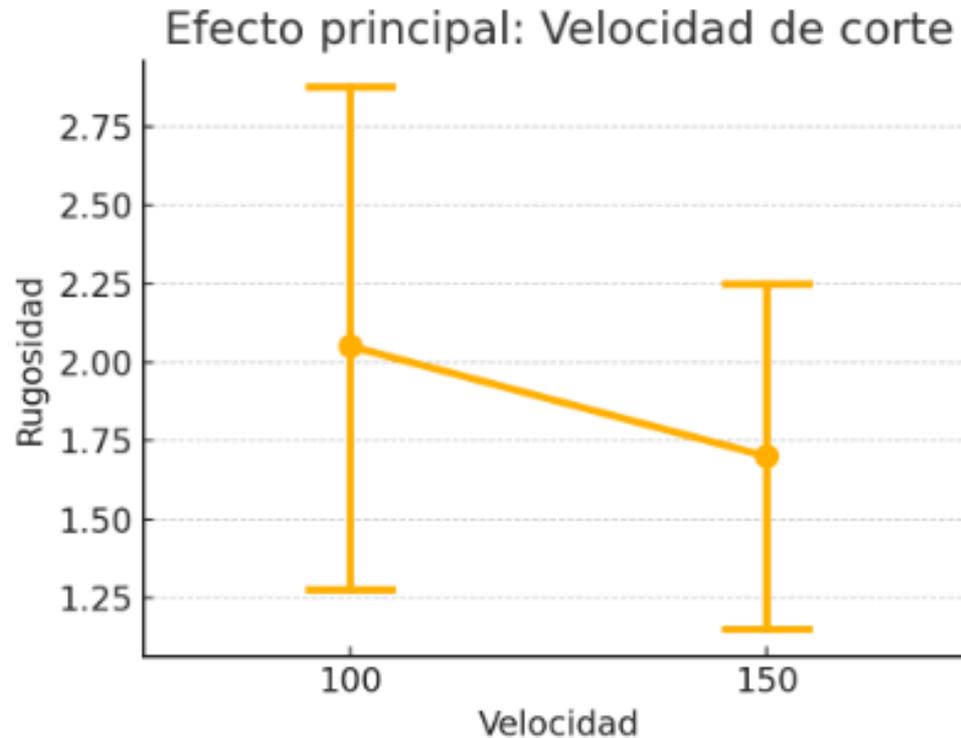
#### PASO 6:

#### Optimización práctica: (La mejor combinación encontrada)

- Velocidad alta (150 m/min)
  - Avance bajo (0.1 mm/rev)
  - Profundidad baja (1.0 mm)
- Predicción: Rugosidad  $R_a \approx 1.0 \mu\text{m}$



# Efectos principales



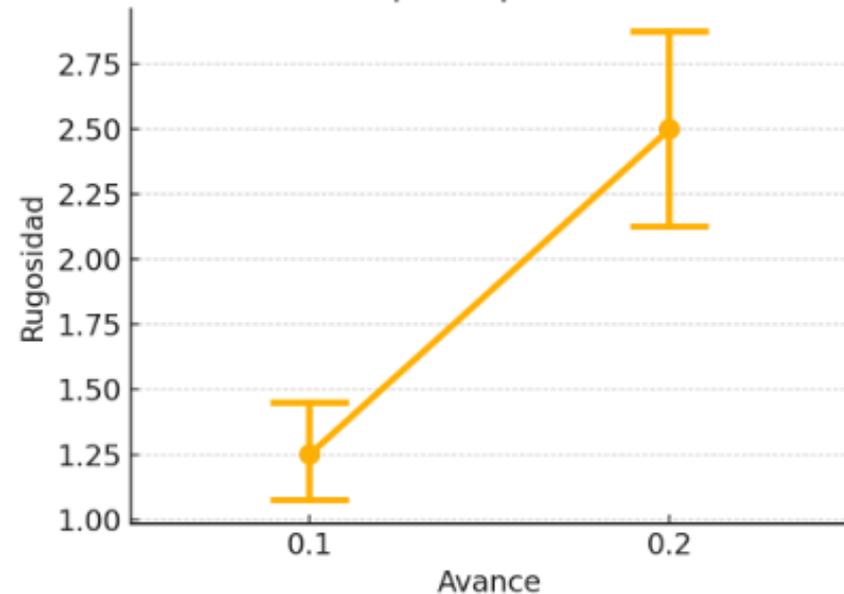
Al aumentar la velocidad de 100  $\rightarrow$  150 m/min, la rugosidad disminuye.

**Interpretación:** una mayor velocidad de corte mejora la calidad superficial, ya que la herramienta genera un acabado más fino.

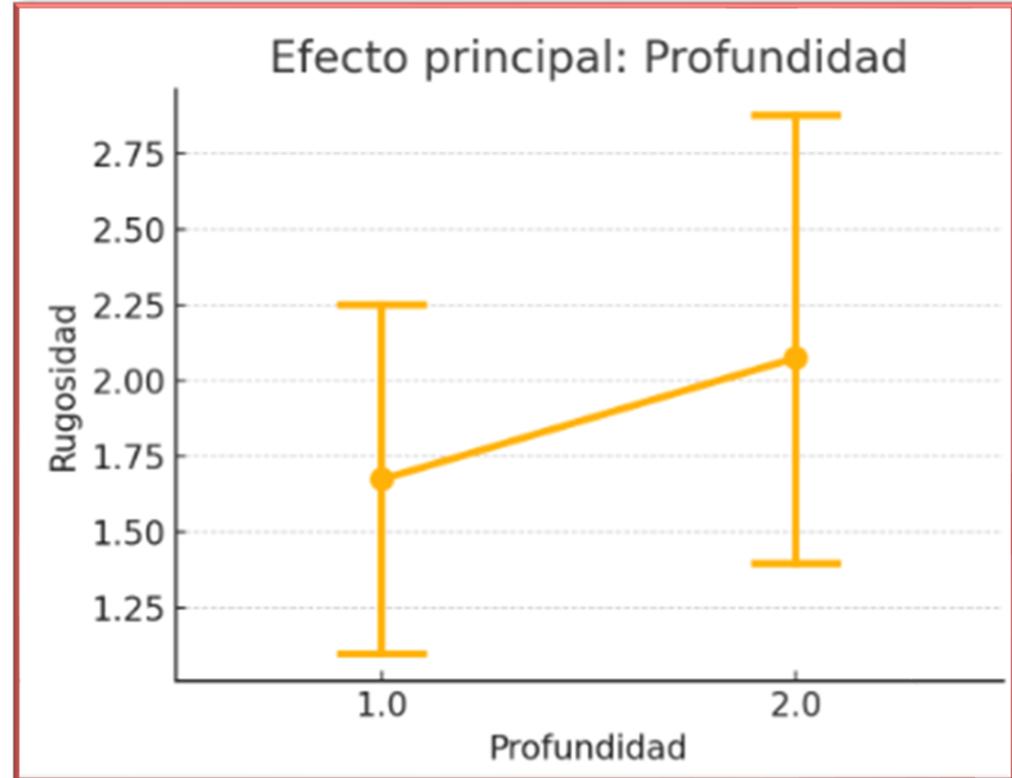
**Conclusión:** velocidad alta es favorable para obtener superficies más lisas.

# Efectos principales

Efecto principal: Avance



Efecto principal: Profundidad

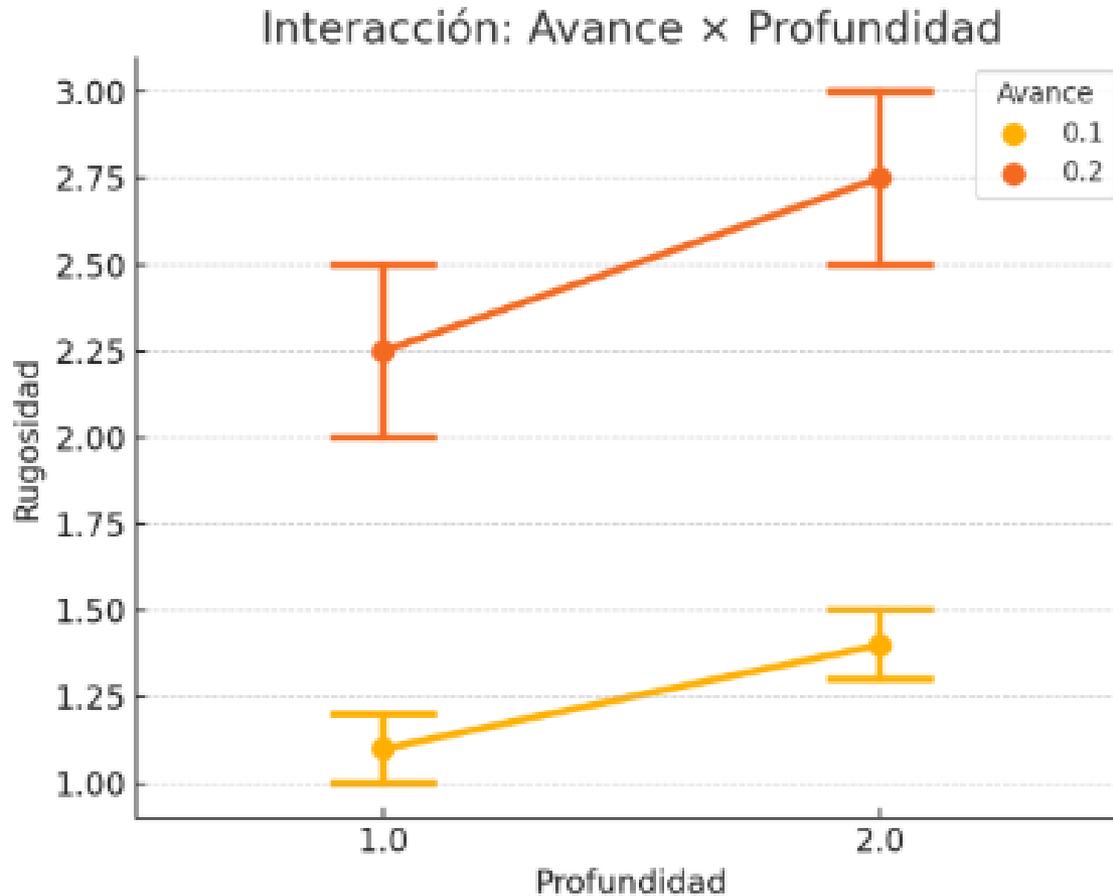


Se observa el efecto más pronunciado: al pasar de 0.1  $\rightarrow$  0.2 mm/rev, la rugosidad aumenta drásticamente.

**Interpretación:** un mayor avance deja “marcas más profundas” en la pieza.

**Conclusión:** el avance es el factor más influyente y debe mantenerse bajo si se busca minimizar rugosidad.

# Interacciones

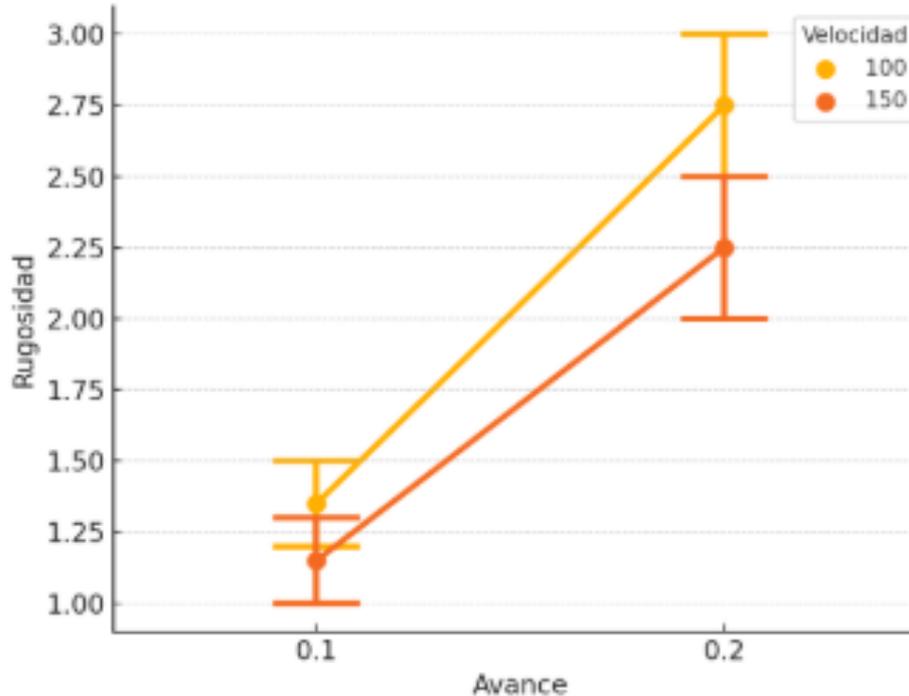


Con avance bajo, la profundidad no cambia mucho la rugosidad. Con avance alto, la rugosidad se dispara si además hay profundidad alta.

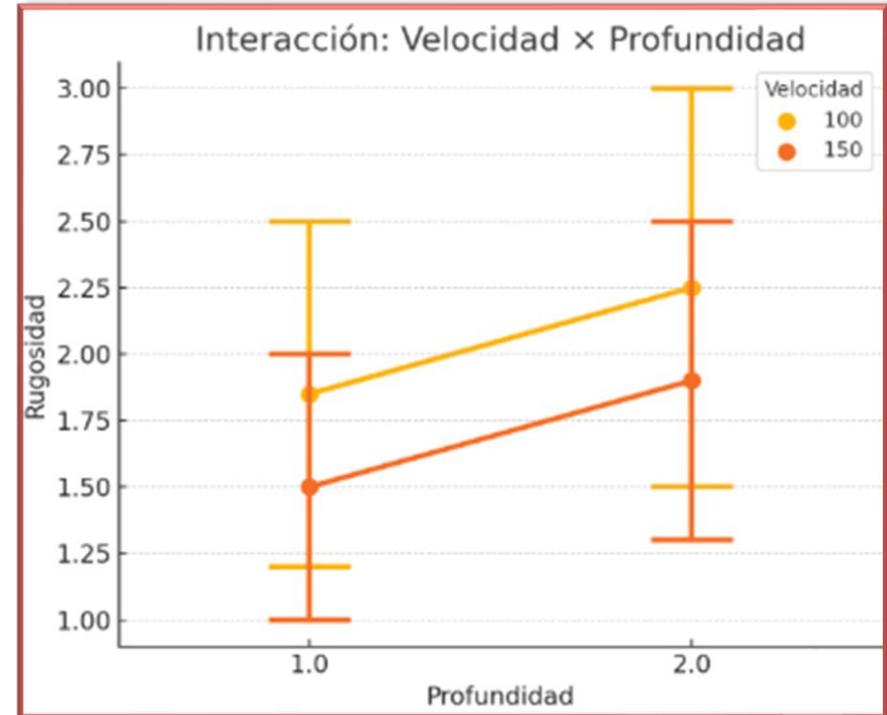
**Conclusión:** el peor escenario es combinar avance alto con profundidad alta.

# Interacciones

Interacción: Velocidad × Avance



Interacción: Velocidad × Profundidad



Con avance bajo (0.1) la velocidad alta realmente mejora el acabado. Sin embargo, con avance alto (0.2) la ventaja de aumentar la velocidad se reduce: la rugosidad sigue siendo elevada.

**Conclusión:** el efecto positivo de la velocidad depende del avance.

# Análisis general

## **El avance es el factor crítico:**

cuanto más bajo, mejor calidad superficial.

La velocidad ayuda, pero su efecto se ve opacado si el avance es muy alto.

La profundidad tiene impacto secundario, pero combinada con avance alto puede generar los peores acabados.

## **Optimización:**

Velocidad alta (150 m/min)

Avance bajo (0.1 mm/rev)

Profundidad baja (1.0 mm)

→ Rugosidad mínima ( $\sim 1.0 \mu\text{m}$ ).

# ANOVA (Análisis de Varianza)

Herramienta estadística central que se usa en el Diseño de Experimentos (DOE).

**RESPONDE A :** ¿Las diferencias que observo en los resultados se deben realmente a los factores que manipulé o son solo producto del azar/ruido experimental?

es el “juez” que te dice si un factor tiene un efecto significativo en la respuesta.

**Relación con (DOE)** Cuando se realiza un diseño de experimento:

Se define factores (ej: temperatura, tiempo).

Se realizan los ensayos y se obtienen respuestas (ej: resistencia).

Se aplica ANOVA para separar y cuantificar (**Variación debida a cada factor**; debida a la **interacción entre factores** y debida al **error experimental**).

Si la variación explicada por un factor es mucho mayor que la variación del error, entonces ese factor influye significativamente.

## Se puede utilizar cuando:

✓ Se comparan medias de más de dos grupos independientes.

**Ejemplo:** comparar el rendimiento académico de alumnos en 3 métodos de enseñanza distintos.

✓ La variable dependiente es cuantitativa (numérica continua).

**Ejemplo:** nivel de glucosa en sangre, tiempo de reacción, productividad, peso, etc.

✓ La variable independiente es categórica (factor) con 2 o más niveles.

**Ejemplo:** tipo de dieta (vegetariana, mediterránea, cetogénica).

## Se puede utilizar cuando:

✓ Se comparan medias de más de dos grupos independientes.

**Ejemplo:** comparar el rendimiento académico de alumnos en 3 métodos de enseñanza distintos.

✓ La variable dependiente es cuantitativa (numérica continua).

**Ejemplo:** nivel de glucosa en sangre, tiempo de reacción, productividad, peso, etc.

✓ La variable independiente es categórica (factor) con 2 o más niveles.

**Ejemplo:** tipo de dieta (vegetariana, mediterránea, cetogénica).

# ANOVA (ejemplo)

Se desea estudiar el efecto de la temperatura y el tiempo de tratamiento sobre la resistencia de una aleación metálica.

Se considera un diseño factorial  $2 \times 2$  con dos repeticiones por combinación de factores:

Temperatura: 200°C y 250°C

Tiempo: 1 h y 2 h

Variable respuesta: resistencia (MPa)

Los datos obtenidos son los siguientes:

Temperatura	Tiempo	Resistencia (MPa)
200	1	150
200	1	152
200	2	160
200	2	158
250	1	170
250	1	172
250	2	180
250	2	182



**Se pide:**

1. Calcular los **promedios por factor** y por combinación de factores.
2. Determinar los **efectos principales** de temperatura y tiempo, y el **efecto de interacción**.
3. Realizar un **ANOVA completo**, incluyendo sumas de cuadrados, grados de libertad, cuadrados medios y valores F.
4. Analizar la **significancia de los factores** y de la interacción.
5. Interpretar los resultados y redactar una **conclusión práctica** sobre qué factor influye más en la resistencia de la aleación.

Calcular los **promedios por factor** y por combinación de factores

Temperatura	Tiempo	Resistencia (MPa)
200	1	150
200	1	152
200	2	160
200	2	158
250	1	170
250	1	172
250	2	180
250	2	182

**Promedio general**

$$\bar{Y} = \frac{150 + 152 + 160 + 158 + 170 + 172 + 180 + 182}{8} = \frac{1,324}{8} = 165.5$$

**Promedio por factor (temperatura)**

$$\bar{Y}_{200} = 155, \quad \bar{Y}_{250} = 176$$

**Promedio por factor (tiempo)**

$$\bar{Y}_{1h} = 161, \quad \bar{Y}_{2h} = 170$$

Determinar los **efectos principales** de temperatura y tiempo, y el **efecto de interacción**.

**Suma de cuadrados total (SST)**

$$SST = \sum (Y_{ij} - \bar{Y})^2$$

Resistencia (
150
152
160
158
170
172
180
182

Calculamos para cada observación:

1.  $(150 - 165.5)^2 = 240.25$

2.  $(152 - 165.5)^2 = 182.25$

3.  $(160 - 165.5)^2 = 30.25$

4.  $(158 - 165.5)^2 = 56.25$

5.  $(170 - 165.5)^2 = 20.25$

6.  $(172 - 165.5)^2 = 42.25$

7.  $(180 - 165.5)^2 = 210.25$

8.  $(182 - 165.5)^2 = 272.25$

$$SST = 240.25 + 182.25 + 30.25 + 56.25 + 20.25 + 42.25 + 210.25 + 272.25 = 1,053.0$$

Determinar los **efectos principales** de temperatura y tiempo, y el **efecto de interacción**.

**Suma de cuadrados por factor (efectos principales)**

$$SSA = r \cdot n_B \sum (\bar{Y}_{A_i} - \bar{Y})^2$$

$r$  = repeticiones por combinación = 2

Temperatura (factor A):

$n_B$  = niveles de factor B = 2

$$\begin{aligned} SSA &= 2 \cdot 2 \cdot [(155 - 165.5)^2 + (176 - 165.5)^2] = \\ &= 4 \cdot [(-10.5)^2 + 10.5^2] = 4 \cdot [110.25 + 110.25] = 4 \cdot 220.5 = 882 \end{aligned}$$

$$SSB = r \cdot n_A \sum (\bar{Y}_{B_j} - \bar{Y})^2$$

Tiempo (factor B):

•  $n_A = 2$

$$\begin{aligned} SSB &= 2 \cdot 2 \cdot [(161 - 165.5)^2 + (170 - 165.5)^2] = \\ &= 4 \cdot [(-4.5)^2 + 4.5^2] = 4 \cdot [20.25 + 20.25] = 4 \cdot 40.5 = 162 \end{aligned}$$

Determinar los **efectos principales** de temperatura y tiempo, y el **efecto de interacción**.

**Suma de cuadrados de interacción (AB)**  $SSAB = r \sum (\bar{Y}_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y})^2$

Se calcula por combinación

1. 200°C, 1h:  $151 - 155 - 161 + 165.5 = 0.5, (0.5)^2 = 0.25$
2. 200°C, 2h:  $159 - 155 - 170 + 165.5 = -0.5, (-0.5)^2 = 0.25$
3. 250°C, 1h:  $171 - 176 - 161 + 165.5 = -0.5, 0.25$
4. 250°C, 2h:  $181 - 176 - 170 + 165.5 = 0.5, 0.25$

### Promedio de combinación

Se promedia cada par de repeticiones dentro de la misma celda.

- 200°C, 1 h:  $(150 + 152)/2 = 302/2 = 151.0$
- 200°C, 2 h:  $(160 + 158)/2 = 318/2 = 159.0$
- 250°C, 1 h:  $(170 + 172)/2 = 342/2 = 171.0$
- 250°C, 2 h:  $(180 + 182)/2 = 362/2 = 181.0$

Se suma y multiplica por 2

$$SSAB = 2 * (0.25 + 0.25 + 0.25 + 0.25) = 2 * 1 = 2$$

### Suma de cuadrados del error (SSE)

$$SSE = SST - SSA - SSB - SSAB = 1053 - 882 - 162 - 2 = 7$$

## Grados de libertad

Fuente	GL
Temperatura (A)	1
Tiempo (B)	1
Interacción AB	1
Error	4
Total	7

## Cuadro ANOVA

Fuente	SC	GL	CM = SC/GL	F = CM/CM_Error
Temperatura	882	1	882	882/1.75 $\approx$ 504
Tiempo	162	1	162	162/1.75 $\approx$ 92.6
Interacción	2	1	2	2/1.75 $\approx$ 1.14
Error	7	4	1.75	
Total	1053	7		

$$g.l._{error} = g.l._{total} - (g.l._A + g.l._B + g.l._{A \times B})$$

$$g.l._{error} = 7 - (1 + 1 + 1) = 7 - 3 = **4**$$

## Cuadro ANOVA

Fuente	SC	GL	CM = SC/GL	F = CM/CM_Error
Temperatura	882	1	882	$882/1.75 \approx 504$
Tiempo	162	1	162	$162/1.75 \approx 92.6$
Interacción	2	1	2	$2/1.75 \approx 1.14$
Error	7	4	1.75	
Total	1053	7		

$$\text{CM\_Error} = \text{SSE} / \text{GL\_Error} = 7 / 4 = 1.75$$

### Interpretación

**1.Temperatura:**  $F \approx 504 \rightarrow$  **muy significativa**

**2.Tiempo:**  $F \approx 92.6 \rightarrow$  **significativa**

**3.Interacción:**  $F \approx 1.14 \rightarrow$  **no significativa**

la combinación de temperatura y tiempo no produce un efecto adicional importante sobre la resistencia. Es decir, los efectos de temperatura y tiempo son independientes.

Conclusión formal: **ambos factores afectan la resistencia, la temperatura es el factor más influyente y no hay interacción significativa.**

## Qué significa esto para el experimento

**1. Temperatura:** es el factor que más aumenta la resistencia. Subir de 200°C a 250°C aumenta la resistencia promedio de 155 MPa a 176 MPa.

**2. Tiempo:** también aumenta la resistencia, pero menos que la temperatura (de 161 a 170 MPa).

**3. Interacción:** la combinación de temperatura y tiempo no tiene un efecto extra. Es decir, la mejora que da aumentar la temperatura se mantiene prácticamente igual sin importar si el tiempo es 1h o 2h.

# Interpretación del valor de F

- El valor de F se calcula como:

$$F = \frac{\text{Variación causada por el factor}}{\text{Variación del error}}$$

- **Regla práctica:**
  - Si  $F \gg 1 \rightarrow$  el factor es significativo.
  - Si  $F \approx 1 \rightarrow$  el factor no es significativo.

En nuestro caso:

- Temperatura:  $F = 504 \rightarrow$  mucho mayor que 1  $\rightarrow$  **muy significativo**
- Tiempo:  $F = 92.6 \rightarrow$  también significativo
- Interacción:  $F = 1.14 \rightarrow$  cercano a 1  $\rightarrow$  **no significativa**

## Grados de libertad (g.l.) en un ANOVA factorial 2×2

Temperatura (A):

$$\text{Niveles} = 2 \rightarrow \text{g.l.} = 2 - 1 = 1$$

Tiempo (B):

$$\text{Niveles} = 2 \rightarrow \text{g.l.} = 2 - 1 = 1$$

Interacción (A×B):

$$\text{g.l.} = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$$

Total:

$$\text{g.l.} = N - 1 = 8 - 1 = 7$$

Error (residuo):

$$g.l._{error} = g.l._{total} - (g.l._A + g.l._B + g.l._{A \times B})$$

$$g.l._{error} = 7 - (1 + 1 + 1) = 7 - 3 = **4**$$

# Quando utilizar ANOVA en (DOE)

ANOVA es una de las herramientas estadísticas más utilizadas porque permite:

Comparar medias de tratamientos o factores.

Determinar si las diferencias observadas son significativas o solo producto del azar.

Evaluar efectos principales e interacciones entre factores.

## Casos en que sí se usa ANOVA:

Experimentos con variables de respuesta cuantitativas y continuas (ej: dureza, resistencia, rendimiento, tiempo).

Cuando se busca comparar medias de varios tratamientos o factores.

## Diseños clásicos como:

- ✓ Completamente al azar (DCA).
- ✓ Bloques completos al azar (DBCA).
- ✓ Factoriales ( $2^k$ ,  $3^k$ , mixtos).
- ✓ Diseños en parcelas divididas.

# Cuando No utilizar ANOVA en (DOE)

## Casos en que NO se usa ANOVA:

- ✓ Cuando la variable de respuesta es categórica o binaria (sí/no, pasa/no pasa): Allí se usan pruebas como Chi-cuadrado, regresión logística o tablas de contingencia.
- ✓ En experimentos de atributos (ej: defectuoso/no defectuoso): Se aplican planes de muestreo, pruebas de proporciones o modelos de Poisson/Binomial.

Si los supuestos de ANOVA **no se cumplen (normalidad de residuos, homogeneidad de varianzas, independencia)**. Se pueden aplicar:

- ✓ Pruebas no paramétricas (Kruskal-Wallis, Friedman).
- ✓ En estudios con una sola variable de dos grupos, se puede usar una prueba t.